ELSEVIER

Clinical Neurophysiology 119 (2008) 1248–1261

www.elsevier.com/locate/clinph

A comparison of quantitative EEG features for neonatal seizure detection

B.R. Greene[a,*], S. Faul[a], W.P. Marnane[a], G. Lightbody[a], I. Korotchikova[b], G.B. Boylan[b]

[a] *Department of Electrical and Electronic Engineering, University College Cork, Room 1.06 Electrical Engineering Building, College Road, Cork, Ireland*
[b] *Department of Paediatrics and Child Health, University College Cork, Ireland*

Accepted 4 February 2008
Available online 1 April 2008

Abstract

Objective: This study was undertaken to identify the best performing quantitative EEG features for neonatal seizures detection from a test set of 21.
Methods: Each feature was evaluated on 1-min, artefact-free segments of seizure and non-seizure neonatal EEG recordings. The potential utility of each feature for neonatal seizure detection was determined using receiver operating characteristic analysis and repeated measures *t*-tests. A performance estimate of the feature set was obtained using a cross-fold validation and combining all features together into a linear discriminant classifier model.
Results: Significant differences between seizure and non-seizure segments were found in 19 features for 17 patients. The best performing features for this application were the RMS amplitude, the line length and the number of local maxima and minima. An estimate of the patient independent classifier performance yielded a sensitivity of 81.08% and specificity of 82.23%.
Conclusions: The individual performances of 21 quantitative EEG features in detecting electrographic seizure in the neonate were compared and numerically quantified. Combining all features together into a classifier model led to superior performance than that provided by any individual feature taken alone.
Significance: The results documented in this study may provide a reference for the optimum quantitative EEG features to use in developing and enhancing neonatal seizure detection algorithms.

Keywords: EEG; Neonatal seizure; Quantitative EEG; Feature extraction

1. Introduction

Seizures are the most common sign of neurological dysfunction in the newborn and require immediate medical attention. The incidence of clinically apparent neonatal seizures in full-term babies is generally reported as around 3 per 1000, and in very preterm babies, approximately 50 per 1000 (Rennie and Boylan, 2007). However, continuous EEG monitoring studies have shown that the clinical diagnosis of neonatal seizures is unreliable, as many seizures display only subtle clinical signs and are often entirely sub-clinical (Murray et al., 2007; Bye and Flanagan, 1995). There is increasing evidence that neonatal seizures have an adverse effect on neurodevelopmental outcome, and predispose to cognitive, behavioural, or epileptic complications in later life (Levene, 2002; McBride et al., 2000).

The EEG is the only reliable method available for the identification and diagnosis of all neonatal seizures but does require continuous access to special expertise for interpretation. Twenty-four hour expertise is not readily available on most neonatal intensive care units and as a result there is an urgent need to develop an automated system that can accurately detect the presence of seizure in the neonate, facilitating prompt medical intervention.

* Corresponding author. Tel.: +353 21 490 3793; fax: +353 21 427 1698.
E-mail address: barryg@rennes.ucc.ie (B.R. Greene).

1388-2457/$34.00
doi:10.1016/j.clinph.2008.02.001

A number of groups have published methods that attempt to automatically detect neonatal seizures (Gotman et al., 1997; Navakatikyan et al., 2006; Greene et al., 2007b,a; Liu et al., 1992; Celka and Colditz, 2002; Aarabi et al., 2006), employing linear and nonlinear signal analysis coupled with a variety of pattern recognition methodologies. Faul et al. (2005a) compared the performance of three well-known neonatal seizure detection algorithms (Gotman et al., 1997; Celka and Colditz, 2002; Liu et al., 1992) and found that none were suitable for clinical use. To date, neonatal seizure detection algorithms have largely resisted transition to widespread clinical use, due to the enormous intra and inter patient variability of the EEG.

In this study, we will examine a number of features, commonly used in automated neonatal seizure detection algorithms and determine which ones best distinguish between seizure and non-seizure EEG. In this way, we aim to recommend the most useful features for deployment in the future as part of the improved seizure detection systems

The dataset and methodology employed in this study were not intended to reflect the performance of each quantitative EEG measure under real-world conditions. Fig. 1 shows the stages involved in an automated seizure detection system; the acquisition and pre-processing of the EEG signal (including artefact identification and rejection), extraction of relevant features from the EEG and finally classification of these features as seizure or non-seizure. This study considers the extraction of optimum features or quantitative EEG measures for stage 2 of this system.

2. Method

2.1. Subjects

A group of 68 neonates out of approximately 14,000 live births, admitted to neonatal intensive care units at the Unified Maternity and Neonatal Services in Cork from May 2004 to May 2006, were recruited into a prospective study of continuous video EEG monitoring in hypoxic ischaemic encephalopathy (HIE). Babies were recruited as soon as possible after birth if they fulfilled 2 or more of the following criteria: initial capillary or arterial pH < 7.1, Apgar score <5 at 5 min, initial capillary or arterial lactate >7 mmol/l or abnormal neurology/clinical seizures. Neonates who did not fulfil these criteria at birth, but later developed clinical seizures were also recruited immediately following their first clinical seizure. The study was approved by the Ethical Committee of Cork University Hospital. Written informed parental consent was obtained for each participant.

2.2. Dataset

The dataset for this study consisted of continuous multi-channel EEG recordings from 17 full-term neonates with seizures, selected from a wider study of 68 neonates with Hypoxic ischaemic encephalopathy (HIE).

Nine channels of EEG (F3, F4, Cz, C3, C4, O1, O2, T3, T4,) and one ECG signal were recorded using the Viasys NicOne video EEG system and sampled at 256 Hz. Electrographic seizures were identified and annotated by two neonatal electroencephalographers (G.B.B. and I.K.) using a referential montage and a mastoid reference electrode.

A number (ordinarily 3) of one minute multi-channel EEG seizure segments were then identified for each patient. Only those channels, noted by the electroencephalographer to have clear electrographic seizure manifestation (referred to as 'involved'), were included in subsequent analysis. For any seizure labelled as "generalised", all EEG channels for that seizure were included in the analysis. Fifty separate seizures, each employing between one and nine channels, were included in the analysis. The same number of non-seizure, multi-channel EEG segments were also selected for each patient, either one hour before seizures began or after an hour of seizure free activity but always on the same postnatal day, all channels from each non-seizure segment were included in the analysis. Table 1 gives the demographic information for each patient used in this study along with the electrographic seizure characteristics of each of the seizures from each of the patients in the dataset.

The most unbiased and realistic method for validating a neonatal seizure detection algorithm would clearly be the use of multi-channel continuous EEG recordings, containing a variety of artefacts and poorly conditioned signals in order to gain a robust estimate of the algorithm's performance under 'real-world' neonatal ICU conditions. Such a dataset however does not easily allow the user to determine the degree of discrimination afforded by each particular

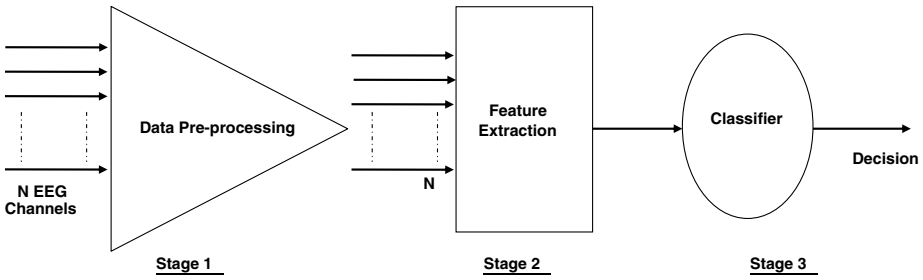

Fig. 1. Schematic of multi-channel EEG based neonatal seizure detection system. This study focuses on stage 2 of this system.

Table 1
Electrographic seizure characteristics for each patient in the dataset

Patient	GA (weeks)	BW (g)	Sex	Seizure #	Seizure location	Frequency (Hz)	Amplitude (μV)
1	39	3030	M	1	O2	1.1–1.3	100
				2	O2	1.1–1.4	90–100
				3	O2	0.8–1.5	80–105
2	40	3700	M	1	C3	1.1–1.3	70–110
				2	C3	1.4–1.9	80–160
3	42	3625	F	1	Generalised	1–1.8	100–300
				2	F4, F3	0.6–0.7	330–350
				3	Generalised	0.8–1.1	80–430
4	40	3670	M	1	F4, C4	1.2–1.5	50–70
				2	F3, C3, T3	1.1–1.2	80–250
				3	F4, C4	1.1–1.4	70–100
5	42	3300	M	1	C3	1.8	50
				2	C3, T3	1.3–1.5	70–95
				3	C3	1.3–1.6	90–120
6	42	3760	M	1	Generalised	1.0–1.5	100-150
				2	F4, C4, F3, C3, Cz	1.1–1.7	60–120
				3	Cz, C3, F3, F4, C4	1.3–1.9	60–90
7	41	3140	F	1	Generalised	1.2–1.4	40–60
				2	Generalised	0.7–1.1	60–120
				3	Generalised	1.0–1.2	55–70
8	41	3750	M	1	C4, T4	1.2	80–100
				2	Generalised	1.1–1.3	55–100
				3	Generalised	1.2–1.4	50–60
9	41	3250	M	1	Generalised	0.6–1.1	100–120
				2	Generalised	0.6–1.0	60–150
				3	C4, T4, O2	0.9–1.6	80–90
10	40	1830	F	1	Generalised	1.4-1.6	30-50
				2	Generalised	1.2–1.5	60–90
				3	C4, T4	1.2–1.9	60–120
11	41	4875	M	1	All except F3 and F4	0.8–1.2	50–130
				2	All except F3, F4	0.5–1.0	100–190
				3	All except F3, F4	0.8–1.1	80–140
12	41	3750	F	1	Generalised	0.8	200–700
				2	Generalised	0.6–1.1	200–650
				3	F4, C4 T4, O2, Cz	0.9–1.0	60–345
13	42	4660	M	1	All except F3 and F4	0.8	50–70
				2	F4, C4, T4, O2	0.8–1.1	80–130
				3	Generalised	0.6–1.1	30–75
14	41	3660	M	1	Generalised	1.2	200-500
				2	Generalised	0.6–1.0	60–190
				3	Generalised	1.3–1.7	70–180
15	40	3650	M	1	Generalised	0.5	100–250
				2	Generalised	0.5–0.7	100–350
				3	Generalised	0.3–0.4	170–280
16	40	4500	M	1	Generalised	1	30–60
				2	Generalised	0.5–1.8	35–135
				3	Generalised	1.6–1.9	25–65
17	40	3100	F	1	Generalised	0.3–0.8	80-200
				2	Generalised	0.6–1.3	80–250
				3	F3,C3, P3, T3, O1	0.4–1.4	60–150

GA is the gestational age of the baby in weeks while BW is the birth weight in grams. 'Seizure #' refers to the characteristics of a specific seizure for each patient. Seizure location refers to the spatial location of the seizure in the EEG segments used. Frequency and amplitude refer to the frequency and amplitude of the seizure discharges as noted by the electroencephalographer.

feature employed in the algorithm. In order to determine which features can best distinguish between non-seizure and seizure neonatal EEG, a sample set of expert labelled artefact-free seizure and non-seizure segments should be employed.

Each 1-min EEG channel segment was band-pass filtered in the range 0.5–34 Hz using a 5th order type-II Chebyshev IIR filter, to remove DC and out of band noise components. The EEG for each 1-min segment, for each channel was considered in 2 s non-overlapping epochs, and all features calculated for each 2 s epoch. The number of epochs included in the analysis for each seizure depended on the number of channels of the segment that were determined to contain seizure activity.

2.3. Features

Twenty-one features in total were analysed for each 2 s EEG epoch. The features employed are derived from linear and nonlinear signal analysis and all have been employed in EEG applications in the past. The features employed in this study can be broadly divided into three separate categories:

1. Frequency domain features.
2. Time domain features.
3. Entropy based features.

All features were implemented using Matlab Version 7 (www.mathworks.com) and are listed in Table 2. Unless otherwise stated all features were implemented by the authors. Third party Matlab software is cited where used.

2.3.1. Frequency domain features

The peak frequency, bandwidth of the dominant spectral peak and spectral power ratio at the dominant frequency features were introduced by Gotman et al. (1997), to distinguish between seizure and non-seizure epochs in the newborn EEG. A periodogram estimate of the power spectral density (PSD) was calculated for each epoch using a 512 point fast Fourier transform (FFT). The dominant frequency was defined to be the frequency in the spectrum with the largest average power in its bandwidth. The bandwidth of the dominant spectral peak was defined as the width in hertz between the two half power points of the dominant spectral peak. The power ratio feature was also discussed by Gotman et al. but was not included in this analysis due to the necessity in the calculation for a moving 'background' EEG, which was not available in this instance.

Throughout the literature a variety of frequency ranges are used in calculating spectral edge frequency (Doyle et al., 2007). Spectral edge frequency (SEF) was calculated according to the methodology discussed by a number of authors (Hudson et al., 1983; Inder et al., 2003; Fell et al., 1996). In this study, we considered three different frequency ranges. The range 2–20 Hz was found to provide the best discrimination between the seizure and non-seizure classes when compared to the ranges 0.5–20 and 0.5–32 Hz. The total EEG power in the range 2–20 Hz was calculated for each 2 s EEG epoch. The SEF was then calculated as that frequency (in Hertz) below which 90% of the total spectral power resides. The total spectral power (TP) was then calculated as the total spectral power in the power spectral density estimate for each epoch in the range 2–20 Hz.

The intensity weighted mean frequency (IWMF) and bandwidth (IWBW) measures were introduced by Evans and McDicken (2000) and discussed by Faul et al. (2005a) in the detection of neonatal seizures. The IWMF is a weighted mean of the frequencies present in the power spectral density estimate for each EEG epoch. IWMF is given by Eq. (1)

$$\mathrm{IWMF} = \frac{\sum_{i=1}^{N_f/2-1} p_i idf}{\sum_{i=1}^{N_f/2-1} p_i}, \tag{1}$$

where i is the frequency bin number, p_i is the estimated spectral power in bin i, $df = f_s/N_f$, where f_s is the sampling frequency and N_f is the total number of frequency bins. The associated bandwidth can be calculated from Eq. (2)

$$\mathrm{IWBW} = \sqrt{\frac{\sum_{i=1}^{N_f/2-1} p_i(\mathrm{IWMF} - idf)^2}{\sum_{i=1}^{N_f/2-1} p_i}}. \tag{2}$$

Wavelet analysis has been previously studied in the context of neonatal seizures (Kitayama et al., 2003) and adult epileptic EEG (Gigola et al., 2004). The EEG for each epoch was decomposed using the Daubechies 4 (Daubechies, 1992) wavelet and decomposed through 8 levels of decomposition. Analysis performed on a separate dataset found that the coefficients of the fifth level of decomposition (the 4th sub-band, corresponding to 1.25–2.5 Hz) are affected most by seizure events and so the fifth level wavelet coefficient was used as the wavelet energy feature for each epoch (Faul, 2007).

2.3.2. Time domain features

Nonlinear energy was used by D'Alessandro et al. (2003) to predict seizures in adult epileptic patients and is calculated for each sample per epoch using Eq. (3), where x is the EEG value for each sample k. The mean nonlinear energy is then taken as a feature for each epoch

$$N(k) = x^2(k) - x(k-1)x(k+1). \tag{3}$$

Esteller et al. (2001) proposed line (curve) length, an estimate of Katz's fractal dimension (Katz, 1988) as a potential feature for epileptic seizure detection in adults. Line length is calculated on an N_s sample epoch basis using Eq. (4)

Table 2
Features employed in this study

	Time domain features
Frequency domain features	
• Bandwidth (BW)	• Line length (L)
• Peak frequency	• RMS amplitude
• Peak power	• Zero crossings
• Spectral edge frequency (SEF)	• Minima and maxima
• Total spectral power (TP)	• Nonlinear energy (N)
• Intensity weighted mean frequency (IWMF)	• Activity (1st Hjorth parameter)
	• Mobility (2nd Hjorth parameter)
• Intensity weighted bandwidth (IWBW)	• Complexity (3rd Hjorth parameter)
• Wavelet energy	• Autoregressive model fit
Entropy based features	
• Shannon entropy ($\mathbf{H_{SH}}$)	• Approximate entropy ($\mathbf{H_{AP}}$)
• Spectral entropy ($\mathbf{H_S}$)	• SVD entropy ($\mathbf{H_{SVD}}$)

$$L = \sum_{k=1}^{N_s} \mathrm{abs}[x(k) - x(k-1)]. \quad (4)$$

The cerebral function monitor (CFM) is perhaps the most widely used tool for the detection and diagnosis of seizure in the neonatal ICU (Murdoch-Eaton et al., 2001; Klebermass et al., 2001; de Vries and Hellstrom-Westas, 2005). The CFM displays a trace of 'Amplitude integrated EEG'. This is calculated from a single EEG channel that has been band-pass filtered between 2 and 15 Hz, with amplification of the higher frequency components, rectified, then applied to a part-linear, part-logarithmic amplitude compression, before being peak-smoothed and then compressed in time (Maynard et al., 1969). Root mean squared EEG amplitude (RMS Amp) represents an estimate of the CFM output and is calculated using Eq. (5)

$$A = \sqrt{\frac{1}{N_s}\sum_{k=1}^{N_s} x^2(k)}. \quad (5)$$

The number of local maxima and minima in an epoch of EEG was discussed by van Putten et al. (2005) and is based on the sum of the number of local minima and maxima for each epoch. It is calculated by counting the number of times in an epoch of EEG that the first derivative of the EEG signal is smaller than a given ε (with $\varepsilon = 0.01$ V).

Similarly the number of 'zero crossings' in the EEG is thought to change during seizure activity (Le Van Quyen et al., 1999; van Putten et al., 2005). The number of positive zero crossings for each epoch is calculated as the sum of all positive zero crossings for each epoch of the zero-meaned EEG, where a sample is defined as a positive zero crossing if the identity in Eq. (6) is true for a given ε (where $\varepsilon = 0.01$ μV)

$$x(k) < \varepsilon \quad \text{and } x(k+5) > \varepsilon. \quad (6)$$

Hjorth parameters (Hjorth, 1975) are based on simple statistical calculations on the EEG and have been used extensively in EEG research (Paivinen et al., 2005; Pfurtscheller and Lopes da Silva, 1999). The first Hjorth parameter, known as 'Activity', is the variance $\boldsymbol{\sigma}_{\mathbf{x}}^2$ of the signal amplitude, where $\boldsymbol{\sigma}_{\mathbf{x}}$ is the standard deviation of the EEG for a given epoch.

The second Hjorth parameter is 'Mobility' and is defined as the square root of the ratio of the activity of the first derivative of the signal ($\boldsymbol{\sigma}_{\mathbf{x}}^{2\prime}$) to the activity of the original signal ($\boldsymbol{\sigma}_{\mathbf{x}}^2$). This can be expressed with standard deviations as Eq. (7)

$$\boldsymbol{\delta}_{\mathbf{x}} = \frac{\boldsymbol{\sigma}_{\mathbf{x}}'}{\boldsymbol{\sigma}_{\mathbf{x}}}. \quad (7)$$

The third Hjorth parameter, 'Complexity', is defined as the ratio of mobility of the first derivative of the signal to the mobility of the signal itself (Eq. (8))

$$\boldsymbol{\lambda}_{\mathbf{x}} = \frac{\boldsymbol{\sigma}_{\mathbf{x}}''/\boldsymbol{\sigma}_{\mathbf{x}}'}{\boldsymbol{\sigma}_{\mathbf{x}}'/\boldsymbol{\sigma}_{\mathbf{x}}}. \quad (8)$$

Autoregressive (AR) modelling methods are widely used in EEG signal processing and analysis (Pfurtscheller et al., 1998; Burke et al., 2005). The following dth order Autoregressive (AR) EEG model, with d equal to 7 (Eq. (9)), was employed in this study using the parameters reported in a separate study (Faul et al., 2007)

$$y(k) = \sum_{i=1}^{d} a_i y(k-i) + \eta(k), \quad (9)$$

where $y(k)$ is the output of the model, $y(1)\ldots y(k-1)$ are the previous outputs, d is the model order, a_i are the model parameters and η is Gaussian white noise. The model parameters are obtained by minimising the sum of least-squares criterion for the forward model and for a time-reversed model. This approach is known as the modified covariance method.

Once the parameters of an AR model have been calculated, its ability to fit another dataset can be analysed to validate the model. If the two datasets are not similar, the fit will be poor. Therefore, an AR model trained on a section of non-seizure EEG provides a poor fit when validated on another section of non-seizure EEG. The error between the EEG (tested on the 2nd second of each 2 s EEG epoch) and the model parameters (trained on the first half of each EEG epoch) provides a feature for each 2 s epoch of EEG.

2.3.3. Entropy features

A number of entropy estimates were employed in this study. Recent evidence suggests that seizure activity represents a drop in complexity of the underlying brain dynamics (Lehnertz and Elger, 1998). Entropy is thought to be a measure of EEG signal complexity and so represents a potential feature for seizure detection.

Several authors have used spectral entropy to quantify the behaviour of the EEG during adult epileptic seizure (D'Alessandro et al., 2003; Kopitzki et al., 1998). The EEG spectral entropy ($\mathbf{H}_{\mathrm{S}}$) was calculated for each seizure and non-seizure epoch using Eq. (10), which normalises $\mathbf{H}_{\mathrm{S}}$ to the range 0–1 (Viertio-Oja et al., 2004)

$$\mathbf{H}_{\mathrm{S}}(X) = -\frac{1}{\log N_f}\sum_f P_f(X)\log_{\mathrm{e}} P_f(X), \quad (10)$$

where $P_f(X)$ is an estimate of the probability density function (PDF) for the EEG epoch X, and is calculated by normalising the power spectral density (PSD) estimate with respect to the total spectral power. The PSD is calculated in the range 0.5–30 Hz for each epoch. N_f is the number of frequency components in the PSD estimate. Applying a histogram estimate of the probability density function, $P_h(X)$ to Shannon's channel entropy formula (Shannon, 1946) yields an estimate of the Shannon entropy ($\mathbf{H}_{\mathrm{SH}}$), as given in Eq. (11) (Moddemeijer, 2007)

$$\mathbf{H}_{\mathrm{SH}}(X) = -\sum_f P_h(X)\log_{\mathrm{e}} P_h(X). \quad (11)$$

Approximate entropy ($\mathbf{H}_{AP}$) is a related measure but is based on estimates of the behaviour of nonlinear dynamical components underlying the EEG (Pincus, 1991; Kaplan and Staffin, 2007). Approximate entropy is defined from Eq. (12)

$$\mathbf{H}_{AP} = \phi^{d_E-1}(r) - \phi^{d_E}(r), \tag{12}$$

where

$$\phi^{d_E}(r) = \frac{1}{M - d_E} \sum_{i=1}^{M-d_E} \log(C_i^{d_E}(r)) \tag{13}$$

and

$$C_i^{d_E}(r) = \{z(j)/M - d_E \mid (\|z(i), z(j)\| < r)/M - d_E\}, \tag{14}$$

where $z(i)$ and $z(j)$ are two points on the underlying state space attractor (a mathematical transposition of the data to a higher dimension in an attempt to determine the underlying structure). M and d_E are the number of points in the time series and the embedding dimension, respectively. The parameter r corresponds to the distance within which neighbouring points must lie on the attractor. $\phi^{d_E}(r)$ represents the likelihood that two nearby points on the attractor remain close to each other. The approximate entropy is therefore a measure of the information contained in the neighbouring points diverging on the attractor.

A measure of complexity which uses both the singular value decomposition (SVD) and the Shannon entropy has been proposed by Roberts et al. (1998) and called the SVD entropy. This process is somewhat similar to the Spectral entropy described above, however instead of drawing on the frequency spectrum, the process is carried out on the singular spectrum. The Roberts et al. complexity measure is based on the singular value decomposition. In practical terms the singular values can tell a lot about the nature of the components of a signal, in particular information about quasi-periodic signals in noise. The singular values $S_1 \ldots S_{d_E}$ can be determined by performing SVD on an embedded matrix. The entropy of the singular spectrum is calculated by first normalising the singular values using Eq. (15)

$$\overline{S}_j = S_j \Big/ \sum_i S_i. \tag{15}$$

For $i,j = 1 \ldots d_E$ (where d_E is the embedding dimension, taken to be 20 in this study), the SVD entropy is then calculated using Eq. (16)

$$H_{SVD} = \sum_{i=1}^{d_E} \overline{S}_i \log \overline{S}_i \tag{16}$$

2.4. Statistical analysis

A number of methods were employed to quantify the behaviour of each feature under seizure and non-seizure conditions as well as to determine the best performing features for this application.

In order to determine if there are significant differences between the seizure and non-seizure class populations for each feature, for each of the 17 patients, a repeated measures paradigm was employed. A significant difference in the seizure/non-seizure means for a given feature and patient was determined by using a paired samples t-test (Weiss, 1995). From histogram analysis, each feature was assumed to be normally distributed under seizure and non-seizure conditions. A separate normal distribution is assumed for each class. All statistical analyses were performed using Matlab version 7.3 and SPSS version 14.0.

Receiver operating characteristic (ROC) curves can be used to determine the relative performances of the feature set. An ROC curve is a plot of class sensitivity against specificity as a threshold parameter is varied. In this study, the seizure/non-seizure decision threshold is varied in the range 0–1. The area under the ROC curve is equivalent to the Mann–Whitney version of the two sample non-parametric Wilcoxon rank-sum statistic (Zweig and Campbell, 1993). The ROC area (calculated using trapezoidal numerical integration) is an effective method for comparing the performance of different features. A random discrimination will give an area of 0.5 under the curve while perfect discrimination between classes will give unity area under the ROC curve.

In order to estimate the utility of each feature for automatic neonatal seizure detection as well as all 21 features taken together, a linear discriminant (LD) classifier model was used. An LD classifier model finds the linear combination of features that best discriminates among classes by maximising Fishers discriminant ratio (Kuncheva, 2004). An unbiased estimate of the classifiers performance was obtained through the use of cross-fold validation. The individual result for the mth patient was obtained by training the classifier model on all the cases for the $m-1$ other patients and testing on the cases for the mth patient where a case is feature vector arising from a given channel for a given epoch of EEG. By rotating through all m possible combinations of training and test sets and taking the mean of the result, an estimate of the classifier's performance was obtained (Kuncheva, 2004). This test provides an indication of the classifier's ability to generalise from the training set and classify from 'unseen' records. In order to ensure that the data from each patient contributed equally to the training of the classifier model, weighting of the mean and covariance matrices by the number of training cases was implemented as previously reported (Greene et al., 2007b). This procedure ensures that a patient with a large amount of data does not unduly influence the training of the classifier model biasing the results, given that the number of channels with seizure included in the analysis varied from patient to patient. It should be noted that this performance estimate cannot be compared directly with the results for multi-channel EEG, but is intended to illustrate

the relative ability of each EEG feature to discriminate between seizure and non-seizure EEG as well the potential benefit of combining a number of EEG features for seizure detection.

3. Results

Fig. 2 shows an example of a focal electrographic seizure for patient 4 (seizure onset on channel C3), this is included to illustrate the difficulty in visually identifying focal neonatal seizure activity.

Fig. 3 shows an example of the generalised electrographic seizure activity for patient 17. Supplementary Fig. S1 shows an example of the generalised electrographic seizure activity for patient 15.

Table 3 gives the mean values for each feature under seizure and non-seizure conditions. The mean is taken across all patients with seizures. Nineteen features showed significant differences ($p < 0.05$) in their mean values under seizure and non-seizure conditions when analysed using a paired samples t-test.

Fig. 4 shows a histogram of the SEF feature under seizure and non-seizure conditions. This provides a graphical indication of the ability of SEF feature to discriminate between seizure and non-seizure electrographic activity.

Fig. 5 shows the class specific histogram for the spectral entropy feature. This histogram shows that the seizure distribution has a lower mean than the non-seizure distribution, however, there is no clear separation between the classes.

The best performing feature was determined by ROC analysis and also to confirm the results provided in Table 3. The area under the ROC curve was used as an index of discrimination. Fig. 6 shows an ROC curve for all 21 features combined into a linear discriminant classifier model as well as ROC curves for five features taken individually, chosen to provide a representative sample of the features analysed.

It is clear from Fig. 6 that the combined features provide better class discrimination than any individual feature. On an individual basis, the best performing feature was found to be the RMS Amp with an area under the ROC curve of 0.86 and the number of maxima and minima, with an area under the ROC curve of 0.85. When all the features were combined and classified in order to obtain an unbiased estimate of their performance on unseen seizure and non-seizure EEG a sensitivity of 81.08% and a specificity of 82.23% were obtained. ROC analysis performed on the dis-

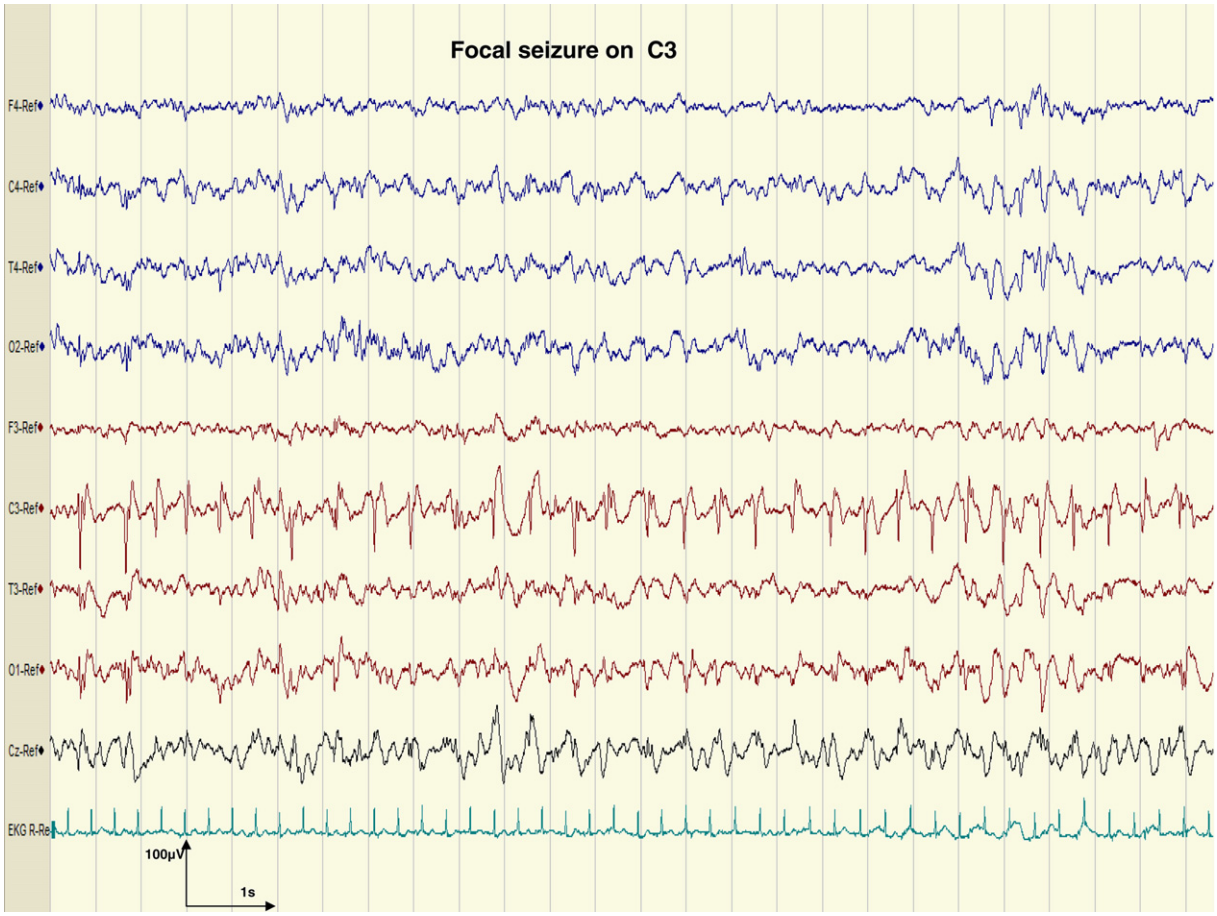

Fig. 2. Multi-channel EEG segment from patient 4 showing a focal, 'non-obvious' electrographic seizure. Seizure is visible on channel C3. Patient 4 had focal, relatively low amplitude seizures. The seizure detection performance for this patient was slightly below average.

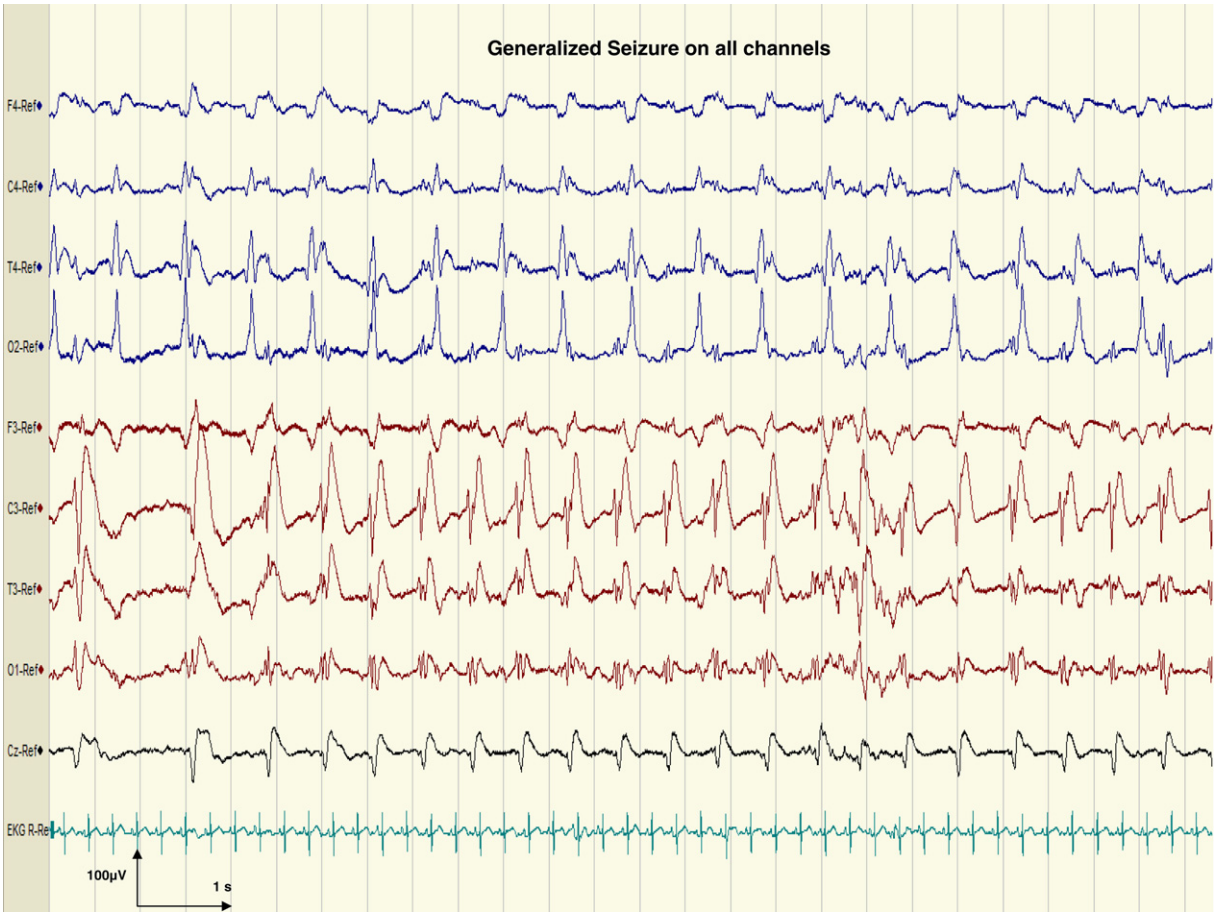

Fig. 3. Example of generalised electrographic seizure activity for patient 17. Seizure is visible on all EEG channels. Patient 17 had relatively low amplitude, generalised seizures. The seizure detection performance for this patient was very poor.

criminant value was found to give an area under the ROC curve of 0.89. Where sensitivity (Sens) is the percentage of 2 s seizure epochs correctly classified as seizure epochs and specificity (Spec) is the percentage of 2 s non-seizure epochs correctly classified as non-seizure. Accuracy (Acc) is the overall percentage of 2 s EEG epochs classified correctly. Subsets of the combined feature set were applied to the classifier to examine the effect of the best and worst performing features on the overall performance. Applying the five best performing features (RMS Amp, maxima and minima, AR fit, line length, total power) to the classifier led to a slight decrease in performance (sensitivity: 80.14 and specificity: 78.96 %). Removing the three worst performing features (complexity, peak freq and IWMF) did not greatly affect performance (sensitivity: 81.16 and specificity: 82.21%).

Table 4 gives a breakdown of the results for all features combined and for each feature classified individually using a linear discriminant classifier model.

Table 5 gives a cross-validation estimate of the performance of all features combined together and classified for each individual patient in the dataset.

Analysis of the seizure detection results for each patient in the data set (given in Table 5) shows that the results for patients 3, 12, 14 and 15 were very good, whereas the results for patients 1, 7, 8 and 17 were relatively poor. The seizures identified for babies 3, 12, 14 and 15 were generalised with high amplitude activity (60–700 μV) (see Supplementary Fig. S2). In contrast the seizures identified for babies 1, 7 and 8 were focal in onset and displayed a lower amplitude seizure activity range (40–100 μV) (see Supplementary Figure S3). Fig. 7 shows the variation of a selection of the features for a given single channel seizure and non-seizure segment for patient 8. Similarly, Fig. 8 illustrates the variation of the same features for the single channel seizure and non-seizure segment for patient 15.

Fig. 7 shows that the seizure EEG for patient 8 is not easily discriminated from the non-seizure EEG using a number of the features considered here. In contrast Fig. 8 shows that the seizure EEG of patient 15 is more easily discriminated from the non-seizure EEG using the selected features. Patient 17 had generalised seizures in the 60–250 μV range. This suggests that high amplitude neonatal seizure activity is more easily detected with conventional features used in seizure detection research than lower amplitude seizure types (which can often be quite similar to the background EEG).

Table 3
Means and standard deviations taken across all patients for each feature under seizure and non-seizure conditions

	Mean		Std. deviation		t	p
	Seizure	Non-seizure	Seizure	Non-seizure		
Frequency domain features						
BW (Hz)	0.74	0.76	0.07	0.04	1.03	$p = 0.3201$
Peak frequency (Hz)	1.16	1.13	0.35	0.59	−0.21	$p = 0.8357$
Peak power (μV^2)	11671.88	1273.73	12353.65	2552.27	−3.60	$p < 0.005$
Spectral edge frequency (Hz)	8.01	11.78	1.39	2.20	5.16	$p < 0.0001$
Total power (μV^2)	4368.11	523.06	3070.32	331.34	−4.85	$p < 0.001$
IWMF (Hz)	2.37	3.59	0.62	1.34	3.12	$p < 0.01$
IWBW (Hz)	2.64	4.21	0.55	1.18	4.34	$p < 0.0005$
Wavelet energy (μV^2)	192.81	79.53	59.42	33.23	−7.10	$p < 0.0001$
Time domain features						
Line length	4.86	2.33	1.52	0.64	−6.22	$p < 0.0001$
RMS Amp (μV)	215.82	63.66	109.71	34.43	−5.17	$p < 0.0005$
Zero crossings	23.31	40.36	7.09	14.37	4.04	$p < 0.001$
Minima and maxima	35.85	123.04	22.38	76.39	4.78	$p < 0.0005$
Nonlinear energy (μV^2)	159.33	16.01	151.08	20.68	−3.79	$p < 0.005$
Activity (μV^2)	75.35	6.83	72.74	7.10	−3.70	$p < 0.005$
Mobility	0.08	0.14	0.02	0.05	3.96	$p < 0.005$
Complexity	4.94	3.85	1.72	1.61	−3.18	$p < 0.01$
AR fit	99.31	98.40	0.23	0.81	−4.31	$p < 0.001$
Entropy based features						
Shannon entropy	3.78	2.97	0.31	0.43	−6.27	$p < 0.0001$
Spectral entropy	0.45	0.60	0.08	0.07	5.37	$p < 0.0005$
Approximate entropy	0.04	0.03	0.00	0.01	−6.01	$p < 0.0001$
SVD entropy	1.60	2.00	0.23	0.26	4.33	$p < 0.001$

The t value and significance level for the t-test for each feature are also given.

4. Discussion

The performance of a set of 21 frequency domain, time domain and entropy based features was examined using a dataset of one minute multi-channel seizure and non-seizure EEG segments from 17 neonates. In this study, we compared the performance of each feature in discriminating seizure activity from non-seizure activity and also to ascertain the range of values occurring for each feature under seizure and non-seizure conditions based on the dataset described above. Paired samples t-tests found that 19 features had significant differences in their means for each patient during seizure when compared to non-seizure EEG. This was confirmed by ROC analysis which found that the best performing feature was RMS Amp followed by the number of local minima and maxima. Applying all features to a linear discriminant classifier model found that superior patient independent seizure detection performance

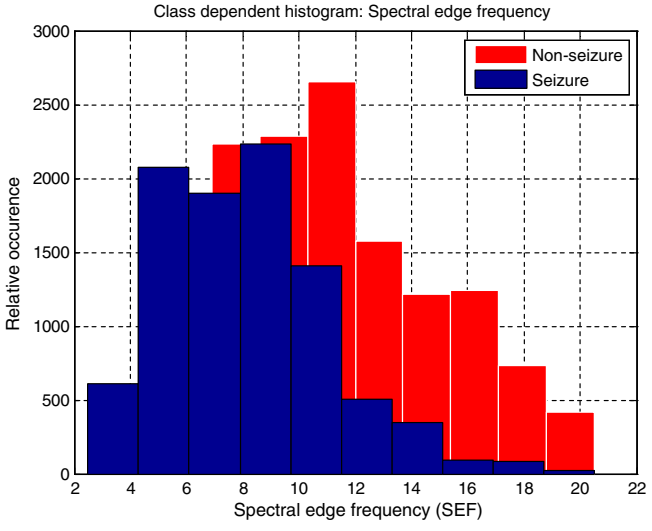

Fig. 4. Class specific histogram for the spectral edge frequency (SEF) feature.

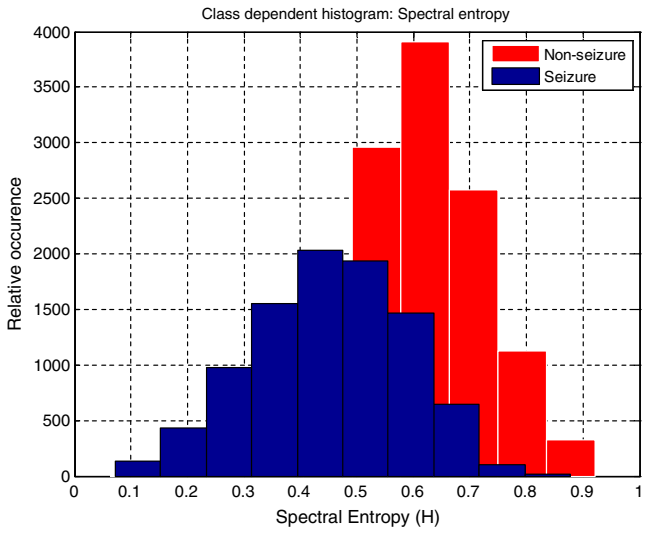

Fig. 5. Class specific histogram for the spectral entropy ($\mathbf{H_S}$) feature.

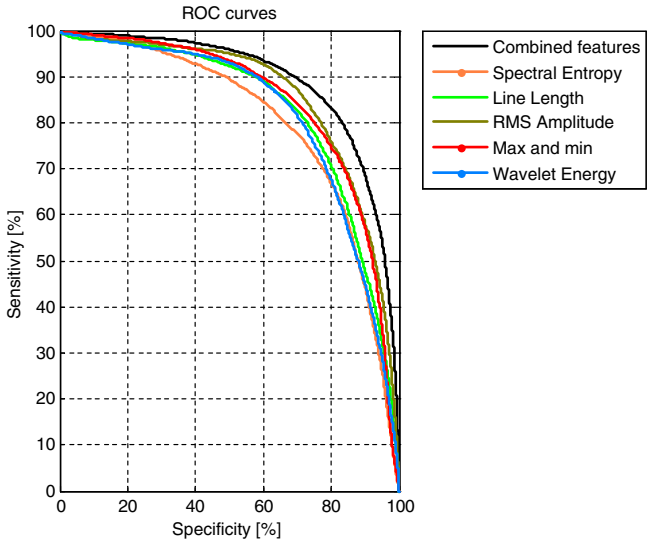

Fig. 6. ROC curves for a number of features. Area under the ROC curve is a good index of seizure/non-seizure class discrimination. Five features from the set used were chosen to provide a representative sample of the features analysed.

might be achieved using a combination of all of the features tested. The reduction in seizure detection performance when small subsets of the feature set were applied to the classifier suggests that even poorly performing features provide some non-redundant seizure specific information.

The features employed in this study arise from both linear and nonlinear signal analysis and have all been employed in EEG studies in the past (Gotman et al., 1997; Esteller et al., 2001; Aarabi et al., 2006; Faul et al., 2005b; Fell et al., 1996; Maiwald et al., 2004; van Putten et al., 2005). Aarabi et al. (2006) discussed the use of a large feature set coupled with a feature selection algorithm based on relevance and redundancy analysis, however they did not provide a discussion of the individual performances of each feature (leading to a 'best performing' feature) or a discussion of the performance of each feature on an artefact-free or 'ideal' set of seizure and non-seizure epochs. Similarly Zarjam et al. (2003) discussed an optimal feature set for neonatal seizure detection, however this feature set contained only features based on wavelet coefficients of seizure and non-seizure EEG segments. van Putten et al. (2005) compared a large number of features from scalp EEG recordings of 16 adult epileptic patients and suggested that no one feature could accurately detect seizures for the patients in their dataset. Similarly in a comparison of a number of linear and nonlinear EEG features on seizures recorded from Sprague–Dawley rats, Paivinen et al. (2005) found that a combination of features performed best in detecting seizures from the EEG.

Table 4
ROC curve area for each feature and all features combined

Feature	Acc (%)	Sens (%)	Spec (%)	ROC
Combined features	81.75	81.08	82.23	0.89
RMS Amp	77.71	64.36	87.21	0.86
Maxima and minima	74.17	87.81	64.46	0.85
AR fit	72.72	88.68	61.36	0.83
Line length	75.81	67.18	81.95	0.83
Total power	73.59	48.33	91.57	0.82
Wavelet energy	74.35	63.16	82.32	0.82
Nonlinear energy	73.99	46.46	93.58	0.81
Activity	73.93	46.43	93.51	0.81
Spectral entropy	74.22	72.69	75.31	0.80
IWBW	71.77	81.09	65.13	0.80
Peak power	72.48	42.42	93.88	0.80
Mobility	71.34	79.52	65.52	0.79
SVD entropy	73.11	73.68	72.71	0.79
Spectral edge frequency	68.90	74.13	65.18	0.77
Shannon entropy	69.66	70.97	68.72	0.76
Zero crossings	69.07	75.86	64.24	0.76
Approximate entropy	68.92	68.06	69.53	0.74
IWMF	66.07	77.63	57.83	0.73
Complexity	66.11	56.22	73.15	0.70
Peak frequency	54.06	38.02	65.48	0.60
Bandwidth	48.84	64.98	37.35	0.51

Results were obtained using a cross-fold validation estimate of the performance of each feature in a patient independent classifier using a linear discriminant classifier model.

Table 5
Patient independent performance estimates for each patient in the dataset

Patient	Acc (%)	Sens (%)	Spec (%)
1	61.63	78.89	59.61
2	68.35	100.00	64.79
3	97.24	97.12	97.34
4	74.04	86.19	70.77
5	70.45	99.15	66.25
6	83.49	82.98	83.85
7	68.93	62.97	73.05
8	58.87	52.34	63.73
9	89.92	85.87	93.16
10	96.30	97.53	95.04
11	82.25	79.49	84.45
12	97.11	94.06	99.75
13	77.01	71.93	81.77
14	84.31	74.44	94.35
15	99.56	99.26	99.87
16	89.59	80.91	99.69
17	67.34	67.13	67.51

Results were obtained using a cross-fold validation estimate of the performance for each patient using a patient independent linear discriminant classifier model.

Given that artefacts are the most well documented confounding factor in biomedical signal algorithms, a clinically useful neonatal seizure detection system must contain a robust and reliable method for the detection and removal of artefacts. This can be attributed to the wide variety of artefacts arising in the EEG in general (Durka et al., 2003) and the neonatal EEG in particular (Gotman et al., 1997; Mizrahi et al., 2004).

The performance of the RMS amplitude feature suggests that this may be one of the best performing features for this application and that under ideal artefact-free conditions; the CFM (which is based on the RMS amplitude feature) maybe a useful tool for the detection of high amplitude generalised neonatal seizures. The performance of the spectral entropy and wavelet energy features provides validation of the seizure detection methods employing entropy measures (D'Alessandro et al., 2003; Greene et al., 2007a)

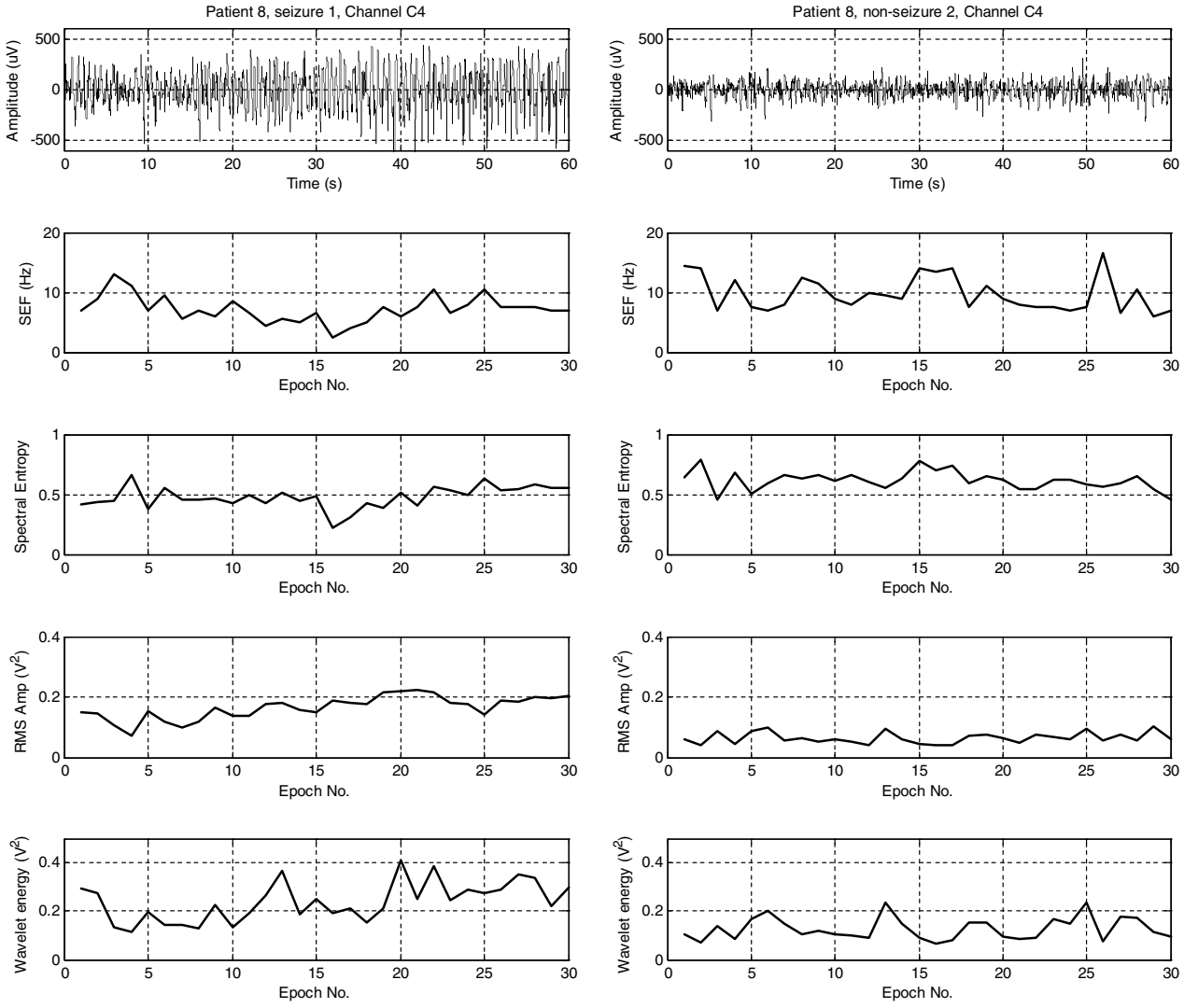

Fig. 7. Variation of SEF, H, RMS amplitude and wavelet energy with seizure and non-seizure activity for patient 8. EEG shown is seizure 1, channel C4 and non-seizure segment 2, channel C4. Patient 8 yielded the poorest seizure detection performance on an individual basis.

and wavelet based measures (Gigola et al., 2004; Nagasubramanian et al., 1997; Zarjam et al., 2003) and suggests that the most informative signal characteristic from a seizure detection point of view are signal complexity metrics as well as heightened energy in 'seizure-specific' frequency bands. Similarly the performance of the line length feature quantifies the visually apparent differences in morphology between seizure and non-seizure neonatal EEG.

In order to obtain a realistic and unbiased estimate of the generalised neonatal seizure detection performance of an algorithm one must use a large number of continuous EEG recordings, containing a variety of artefacts and poorly conditioned signals in order to gain a robust estimate of the algorithm's performance under 'real-world' neonatal ICU conditions. However, such a validation philosophy does not allow accurate determination of how well a given feature discriminates pure 'unblemished' EEG seizure activity from non-seizure EEG activity. In order to determine which EEG features can best distinguish between the (sometimes very similar) abnormal non-seizure EEG and seizure EEG one must use a set of sample or *prototypical*, artefact-free seizure and non-seizure segments, each of which has been labelled as such by an expert in neonatal EEG. A cross-fold validation estimate of the performance of each of the features as well as all the features combined, as part of a patient independent neonatal seizure detection system was included for completeness. It must be emphasized that this performance estimate does not reflect the performance of the system on continuous multi-channel 'raw' EEG.

The issue of redundant information is one that must be addressed when considering quantitative EEG features. Many of the features considered in this study are essentially measuring the same property of the seizure EEG. For example the total power, nonlinear energy, wavelet energy and RMS amplitude features are all characterising the power in the EEG signal, as a result a number of the features are highly correlated with each other. Furthermore, some nonlinear measures, when compared to linear measures have

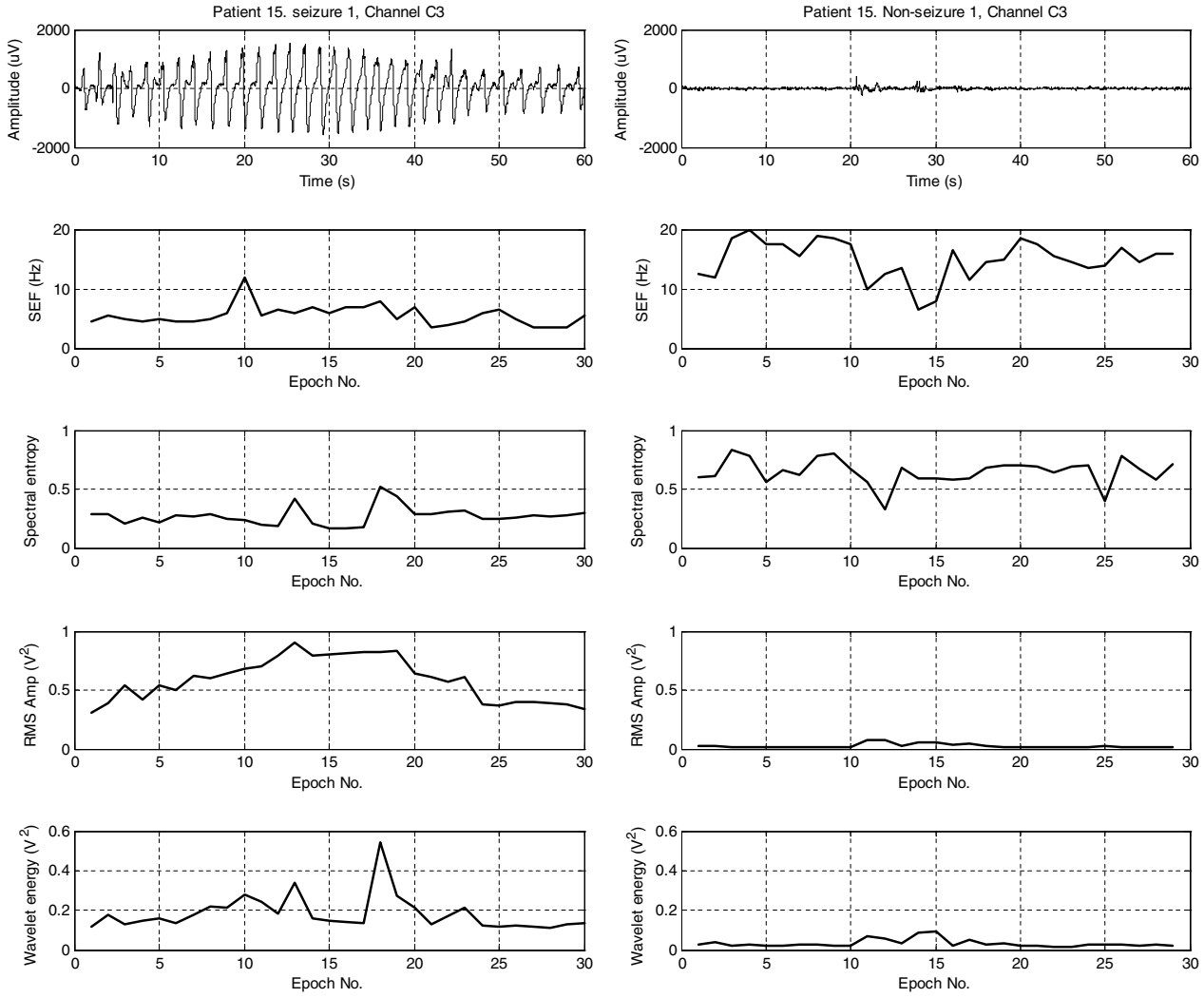

Fig. 8. Variation of SEF, H, RMS amplitude and wavelet energy with seizure and non-seizure activity for patient 15. EEG shown is seizure 1, channel C3 and non-seizure segment 1, channel C3. Patient 15 yielded the best seizure detection performance on an individual basis.

been hypothesised to be a different means of examining the same signal property (Martinerie et al., 2003).

A number of electrographic seizure types were included in this analysis including focal, multi-focal, and generalised seizures. However all seizures were assigned the categorical label of 'Seizure', so the question of how individual features perform in detecting differing seizure types is not addressed. Future work will consider the variation of a number of quantitative features with different gestational ages, pathologies and seizure types. Similarly, given that the EEG is highly age dependent, future work will examine the effect of gestational age on the seizure and non-seizure characteristics of each feature.

5. Conclusion

The individual performances of 21 quantitative EEG features for detecting electrographic seizure in the neonate were compared and numerically quantified in order to determine the best features to be used in a neonatal seizure detection algorithm. Combining all features together into a statistical classifier model, yielded superior seizure detection performance, when compared to the performance yielded by any feature taken alone. On an individual basis the RMS amplitude, line length and the number of local minima and maxima were found to give the best seizure detection performance. It is anticipated that the results from this study will provide a platform from which to develop and enhance neonatal seizure detection algorithms.

Acknowledgement

This project was funded by a Science Foundation Ireland (SFI/05/PICA/1836) principal investigator career advancement award.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.clinph.2008.02.001.

References

Aarabi A, Wallois F, Grebe R. Automated neonatal seizure detection: a multistage classification system through feature selection based on relevance and redundancy analysis. Clin Neurophysiol 2006;117:328–40.

Burke DP, Kelly SP, de Chazal P, Reilly RB, Finucane C. A parametric feature extraction and classification strategy for brain–computer interfacing. IEEE Trans Rehabil Eng 2005;13:12–7.

Bye AME, Flanagan D. Spatial and temporal characteristics of neonatal seizures. Epilepsia 1995;36:1009–16.

Celka P, Colditz P. A computer-aided detection of EEG seizures in infants: a singular-spectrum approach and performance comparison. IEEE Trans Biomed Eng 2002;49:455–62.

D'Alessandro M, Esteller R, Vachtsevanos G, Hinson A, Echauz J, Litt B. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. IEEE Trans Biomed Eng 2003;50:603–15.

Daubechies I. Ten lectures on wavelets. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1992.

de Vries LS, Hellstrom-Westas L. Role of cerebral function monitoring in the newborn. Arch Dis Child Fetal Neonatal Ed 2005;90:F201–7.

Doyle OM, Greene BR, Murray DM, Marnane L, Lightbody G, Boylan GB. The effect of frequency band on quantitative EEG measures in neonates with Hypoxic-ischaemic encephalopathy. In: Proceedings of the 29th international conference of the IEEE-EMBS, 2007. p. 717–21.

Durka PJ, Klekowicz H, Blinowska KJ, Szelenberger W, Niemcewicz S. A simple system for detection of EEG artifacts in polysomnographic recordings. IEEE Trans Biomed Eng 2003;50:526–8.

Esteller R, Echauz J, Tcheng T, Litt B, Pless B. Line length: an efficient feature for seizure onset detection. In: Proceedings of the 23rd annual international conference of the IEEE engineering in medicine and biology society, vol. 2, 2001. p. 1707–10.

Evans DH, McDicken WN. Doppler ultrasound. Physics, instrumentation and signal processing. Wiley; 2000.

Faul S. Automated neonatal seizure detection. Ph.D. University College Cork, 2007.

Faul S, Boylan G, Connolly S, Marnane L, Lightbody G. An evaluation of automated neonatal seizure detection methods. Clin Neurophysiol 2005a;116:1533–41.

Faul S, Boylan GB, Connolly S, Marnane S, Lightbody G. Chaos theory analysis of the newborn EEG: is it Worth the Wait? In: IEEE international workshop on intelligent signal processing (WISP'05), 2005. p 381–6.

Faul S, Gregorcic G, Boylan G, Marnane L, Lightbody G, Connolly S. Gaussian process modelling of EEG for the detection of neonatal seizures. IEEE Trans Biomed Eng 2007;54:2151–62.

Fell J, Roschke J, Mann K, Schaffner C. Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. Electroencephalogr Clin Neurophysiol 1996;98:401–10.

Gigola S, Ortiz F, D'Attellis CE, Silva W, Kochen S. Prediction of epileptic seizures using accumulated energy in a multiresolution framework. J Neurosci Methods 2004;138:107–11.

Gotman J, Flanagan D, Zhang J, Rosenblatt B, Bye A, Mizrahi EM. Automatic seizure detection in newborns: methods and initial evaluation. Electroencephalogr Clin Neurophysiol 1997;103:356–62.

Greene BR, Boylan GB, Reilly RB, de Chazal P, Connolly S. Combination of EEG and ECG for improved automatic neonatal seizure detection. Clin Neurophysiol 2007a;118:1348–59.

Greene BR, de Chazal P, Boylan GB, Connolly S, Reilly RB. Electrocardiogram based neonatal seizure detection. IEEE Trans Biomed Eng 2007b;54:673–82.

Hjorth B. An on-line transformation of EEG scalp potentials into orthogonal source derivations. Electroencephalogr Clin Neurophysiol 1975;39:526–30.

Hudson RJ, Stanski DR, Saidman LJ, Meathe E. A model for studying depth of anesthesia and acute tolerance to thiopental. Anesthesiology 1983;59:301–8.

Inder TE, Buckland L, Williams CE, Spencer C, Gunning MI, Darlow BA, et al. Lowered electroencephalographic spectral edge frequency predicts the presence of cerebral white matter injury in premature infants. Pediatrics 2003;111:27–33.

Kaplan D, Staffin P. HRV software, 2007. Available from: http://www.macalester.edu/~kaplan/hrv/doc/.

Katz M. Fractals and the analysis of waveforms. Comp Biol Med 1988;18:145–56.

Kitayama M, Otsubo H, Parvez S, Lodha A, Ying E, Parvez B, et al. Wavelet analysis for neonatal electroencephalographic seizures. Pediatr Neurol 2003;29:326–33.

Klebermass K, Kuhle S, Kohlhauser-Vollmuth C, Pollak A, Weninger M. Evaluation of the cerebral function monitor as a tool for neurophysiological surveillance in neonatal intensive care patients. Childs Nerv Syst 2001;17:544–50.

Kopitzki K, Warnke PC, Timmer J. Quantitative analysis by renormalized entropy of invasive electroencephalograph recordings in focal epilepsy. Phys Rev E 1998;58:4859–64.

Kuncheva LI. Combining pattern classifiers; methods and algorithms. Wiley; 2004.

Le Van Quyen M, Martinerie J, Baulac M, Varela F. Anticipating epileptic seizures in real time by a non-linear analysis of similarity between EEG recordings. Neuroreport 1999;10:2149–55.

Lehnertz K, Elger CE. Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity. Phys Rev Lett 1998;80:5019–22.

Levene M. The clinical conundrum of neonatal seizures. Arch Dis Child 2002;86:75–7.

Liu A, Hahn JS, Heldt GP, Coen RW. Detection of neonatal seizures through computerized EEG analysis. Electroencephalogr Clin Neurophysiol 1992;82:30–7.

Maiwald T, Winterhalder M, Aschenbrenner-Scheibe R, Voss HU, Schulze-Bonhage A, Timmer J. Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic. Physica D Nonlinear Phenomena 2004;194:357–68.

Martinerie J, Le Van Quyen M, Baulac M, Renault B. Reply to "Prediction of epileptic seizures: are nonlinear methods relevant?". Nat Med 2003;9:242.

Maynard D, Prior PF, Scott DF. Device for continuous monitoring of cerebral activity in resuscitated patients. Br Med J 1969;4:545–6.

McBride MC, Laroia N, Guillet R. Electrographic seizures in neonates correlate with poor neurodevelopmental outcome. Neurology 2000;55:506–14.

Mizrahi EM, Hrachovy RA, Kellaway P. Atlas of neonatal electroencephalography. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.

Moddemeijer R. Matlab library, 2007. Available from: http://www.cs.rug.nl/~rudy/matlab/.

Murdoch-Eaton D, Darowski M, Livingston J. Cerebral function monitoring in paediatric intensive care: useful features for predicting outcome. Dev Med Child Neurol 2001;43:91–6.

Murray DM, Boylan GB, Ali I, Ryan CA, Murphy BP, Connolly S. Defining the gap between electrographic seizure burden, clinical expression, and staff recognition of neonatal seizures. Arch Dis Child Fetal Neonatal Ed 2007, adc.2005.086314.

Nagasubramanian S, Onaral B, Clancy R. On-line neonatal seizure detection based on multi-scale analysis of EEG using wavelets as a tool. In: Proceedings of the 19th annual international conference of the

IEEE engineering in medicine and biology society, vol. 3, 1997. p. 1289–92.

Navakatikyan MA, Colditz PB, Burke CJ, Inder TE, Richmond J, Williams CE. Seizure detection algorithm for neonates based on wave-sequence analysis. Clin Neurophysiol 2006;117:1190–203.

Paivinen N, Lammi S, Pitkanen A, Nissinen J, Penttonen M, Gronfors T. Epileptic seizure detection: a nonlinear viewpoint. Comput Methods Programs Biomed 2005;79:151–9.

Pfurtscheller G, Lopes da Silva FH. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 1999;110:1842–57.

Pfurtscheller G, Neuper C, Schlogl A, Lugger K. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. IEEE Trans Rehabil Eng 1998;6:316–25.

Pincus S. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci 1991;88:2297–301.

Rennie JM, Boylan GB. Treatment of neonatal seizures. Arch Dis Child Fetal Neonatal Ed 2007;92:F148–50.

Roberts SJ, Penny W, Rezek I. Temporal and spatial complexity measures for EEG-based brain–computer interfacing. Med Biol Eng Comput 1998;37:93–9.

Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1946;27:379–423.

van Putten MJAM, Kind T, Visser F, Lagerburg V. Detecting temporal lobe seizures from scalp EEG recordings: a comparison of various features. Clin Neurophysiol 2005;116:2480–9.

Viertio-Oja H, Maja V, Sarkela M, Talja P, Tenkanen N, Tolvanen-Laakso H, et al. Description of the entropy algorithm as applied in the datex-ohmeda entropy module. Acta Anaesthesiol Scand 2004;48:154–61.

Weiss NA. Introductory Statistics. Addison Wesley; 1995.

Zarjam P, Mesbah M, Boashash B. An optimal feature set for seizure detection systems for newborn EEG signals, In: Circuits and systems, 2003. ISCAS'03. Proceedings of the 2003 international symposium on, vol. 5, 2003. p. V-33–6.

Zweig M, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem 1993;39:561–77.